METHODS AND MATERIALS FOR THE PRODUCTION OF SHIKIMIC ACID

[0001] This invention was made with Government support under Contract 08-R1GM065541A, awarded by the National Institutes of Health. The Government may have certain rights in this invention.

INTRODUCTION

[0002] The present invention relates to methods, materials and organisms for the production of shikimic acid and related compounds. In particular, such methods relate to microbial synthetic processes using pyruvate.

[0003] Shikimic acid is an attractive chiral compound useful in a variety of synthetic reaction. It has a highly functionalized, six-membered carbocyclic ring, and multiple asymmetric centers. A metabolic intermediate of aromatic amino acid biosynthesis, shikimic acid is a commercially valuable chiral starting material in the synthesis of neuraminidase inhibitors effective in the treatment of influenza. See, e.g., Kim. C. U. et al., *J. Am. Chem. Soc. 119*:681 (1997); and Rohloff, J. C. et al., *J. Org. Chem. 63*:4545 (1998). Chiral, as well as aromatic chemicals, can also be synthesized from shikimic acid. For example, acid catalyzed dehydration of shikimic acid affords p-hydroxybenzoic acid, which has an annual production of over seven million kilograms, and is the key precursor to parabens and a monomer used in the synthesis of liquid crystal polymers. Shikimic acid has also been used as the starting point for synthesis of a large combinatorial library of molecules. See, e.g., Tan, D. S. et al., *J. Am. Chem. Soc. 120*:8565 (1998).

[0004] Shikimic acid may be obtained via tedious multi-step isolation procedures from plants. Unfortunately, current isolation of shikimic acid from the fruit of *Illicium* plants (Haslem, E., "Shikimic Acid: Metabolism and Metabolites," Wiley & Sons, New York, pp. 40-42 (1993)) precludes its use in kilogram-level synthesis

[0005] Microbial synthesis of shikimic acid is disclosed in U.S. Patent 5,168,056, Frost, issued December 1, 1992; U.S. Patent 5,272,073, Frost et al., issued December 21, 1993: U.S. Patent 5,629,181, Frost et al., issued May 13, 1997; and U.S. Patent 6,613,552, Frost et al., issued September 2, 2003. Such methods employ variations on the common pathway for aromatic amino acid biosynthesis.

[0006] Phosphoenolpyruvate (PEP)is a substrate for the first committed step in the shikimate pathway (Figure 1) and is also used by the carbohydrate phosphotransferase (PTS) system for microbial transport and phosphorylation of glucose. The resulting competition between the shikimate pathway and PTS-mediated glucose transport for cytoplasmic supplies of phosphoenolpyruvate limits the concentrations and yields of natural products microbially synthesized by way of the shikimate pathway.

SUMMARY

[0007] The present invention provides methods for producing shikimic acid comprising enzyme-catalyzed condensation of pyruvate with D-erythrose 4-phosphate to form 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP). Also provided are methods for making shikimic acid, as depicted in Figure 2, and transfected microbes expressing 2-keto-3-deoxy-6-phosphogalactonate (KDPGal) aldolase.

[0008] The present invention further provides:

- Recombinant KDPGal aldolase polypeptides having the ability to catalyze the condensation of pyruvate and E4P to from DAHP, containing at least one mutation that is X10V, X28L or X28M, X42T, X85A, X154F, or X196I; KDPGal aldolase polypeptides containing at least one mutation that is I10V, V28L or V28M, S42T, V85A, V154F, or F196I;
- KDPGal aldolase polypeptides having, apart from one of these mutations, the amino acid sequence of any one of SEQ ID NO:2, SEQ ID NO:4, and SEQ ID NO:6, or an amino acid sequence at least 50% homologous to one of these; recombinant KDPGal aldolase polypeptides whose amino acid sequences are variants of a native KDPGal aldolase amino acid sequence;
- Nucleic acid encoding such a recombinant KDPGal aldolase polypeptide; vectors containing such nucleic acid;
- Enzymatic pathways capable of converting pyruvate and D-erythrose 4-phosphate (E4P) into 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP), by virtue of their having at least one KDPGal aldolase; such enzymatic pathways also capable of converting DAHP to DHQ by virtue of

their having at least one DHQ synthase, and optionally at least one DHQ dehydratase, and further optionally at least one shikimate dehydrogenase;

- Methods for the production of shikimate or a shikimate intermediate, such as DAHP, DHQ, or DHS, by growing a recombinant host cell containing nucleic acid encoding at least one KDPGal aldolase and at least one DHQ synthase, such that the cell expresses those enzymes. Methods for converting pyruvate and E4P to DAHP or a derivative of DAHP in vitro or in vivo/in cyto;
- The use of a recombinant KDPGal aldolase to produce DAHP from pyruvate and E4P; the use of a combination of recombinant KDPGal aldolase and DHQ synthase to produce DHQ.
- Process for preparing recombinant cells capable of expressing a KDPGal aldolase, and thus of converting pyruvate and E4P to DAHP by providing a host cell capable of synthesizing pyruvate and E4P, providing a vector containing a polynucleotide from which said host cell can express a KDPGal aldolase, and transforming said cell with said vector to produce a transformed cell, and, optionally, expressing the KDPGal aldolase, whereupon the cell converts pyruvate and E4P to DAHP.
- Recombinant cells prepared thereby;
- Processes for preparing DAHP or a derivative thereof, by providing (A) a supply of E4P and pyruvate, (B) a KDPGal aldolase, and optionally a DHQ synthase, (C) an aqueous medium; contacting the KDPGal aldolase with the E4P and pyruvate under conditions in which the KDPGal aldolase can catalyze the formation of DAHP therefrom, and optionally contacting the DAHP with the DHQ synthase under conditions in which the DHQ synthase can catalyze the formation of 3-dehydroquinate from the DAHP; and optionally recovering at least one of DAHP, DHQ, DHS, or a further derivative thereof;
- In vivo embodiments of such methods, pathways, and cells further including a recombinant transketolase or a recombinant transaldolase;
- Kits containing a KDPGal aldolase preparation, with instructions for the use thereof to convert pyruvate and E4P to DAHP, and optionally with instructions for the conversion of said DAHP to at least one derivative thereof;

- Kits containing a cell capable of expressing a KDPGal aldolase, with instructions for the use thereof to convert pyruvate and E4P to DAHP, and optionally with instructions for the conversion of said DAHP to at least one derivative thereof;
- Kits containing nucleic acid from which a cell can express at least one KDPGal aldolase, with instructions for the use thereof to transform a cell to produce a transformed cell that is capable of onverting pyruvate and E4P to DAHP, and optionally to at least one derivative thereof.

FIGURES

[0009] Figure 1 depicts a synthetic scheme for 3-dehydroshikimate.

[0010] Figure 2 depicts a synthetic scheme for shikimic acid and intermediate compounds.

[0011] Figure 3 sets forth restriction enzyme maps for plasmids useful in the methods of this invention.

[0012] Figure 4 is a plot depicting showing the growth of organisms among those useful herein.

[0013] Figure 5 depicts a synthetic scheme for enhanced conversion of glucose to DAHP via a pyruvate-based pathway according to the present invention. The numbers are the relative fluxes involved in converting 7 mol of glucose into DAHP. Enzymes are: Pps, PEP synthase; Tkt, transketolase; Tal, transaldolase. Metabolites are: G6P, glucose 6-phosphate; F6P, fructose 6-phosphate; 1,6-FDP, 1,6-fructose diphosphate; DHAP, dihydroxyacetone phosphate; GAP, glyceraldehyde 3-phosphate; R5P: ribose 5-phosphate; X5P, xylulose 5-phosphate; S7P, sedoheptulose 7-phosphate; PYR, pyruvate. This scheme shows that E4P production can be improved by enhancing expression of Tkt and/or Tal, thereby increasing the synthesis of DAHP by KDPGal aldolase.

[0014] It should be noted that the figures set forth herein are intended to exemplify the general characteristics of an apparatus, materials and methods among those of this invention, for the purpose of the description of such embodiments herein. These figures may not precisely reflect the

characteristics of any given embodiment, and are not necessarily intended to define or limit specific embodiments within the scope of this invention.

DESCRIPTION

[0015] The present invention provides methods, materials and organisms for producing shikimic acid and intermediates. The following definitions and non-limiting guidelines must be considered in reviewing the description of this invention set forth herein.

[0016] The headings (such as "Introduction" and "Summary,") and sub-headings used herein are intended only for general organization of topics within the disclosure of the invention, and are not intended to limit the disclosure of the invention or any aspect thereof. In particular, subject matter disclosed in the "Introduction" may include aspects of technology within the scope of the invention, and may not constitute a recitation of prior art. Subject matter disclosed in the "Summary" is not an exhaustive or complete disclosure of the entire scope of the invention or any embodiments thereof.

[0017] The citation of references herein does not constitute an admission that those references are prior art or have any relevance to the patentability of the invention disclosed herein. Any discussion of the content of references cited in the Introduction is intended merely to provide a general summary of assertions made by the authors of the references, and does not constitute an admission as to the accuracy of the content of such references. All references cited in the Description section of this specification are hereby incorporated by reference in their entirety.

[0018] The description and specific examples, while indicating embodiments of the invention, are intended for purposes of illustration only and are not intended to limit the scope of the invention. Moreover, recitation of multiple embodiments having stated features is not intended to exclude other embodiments having additional features, or other embodiments incorporating different combinations the stated of features.

[0019] As used herein, the words "preferred" and "preferably" refer to embodiments of the invention that afford certain benefits, under certain circumstances. However, other embodiments may also be preferred, under the same or other circumstances. Furthermore, the recitation of one or more preferred embodiments does not imply that other embodiments are not

useful, and is not intended to exclude other embodiments from the scope of the invention.

[0020] As used herein, the word "include," and its variants, is intended to be non-limiting, such that recitation of items in a list is not to the exclusion of other like items that may also be useful in the materials, compositions, devices, and methods of this invention.

[0021] The word :"recombinant" is used herein to indicate that nucleic acid manipulation was employed. As a result, phrases such as "recombinant" nucleic acid, "recombinant" polypeptide, and "recombinant" cell to entities that were produced, at least in part, by nucleic acid manipulation.

Sequence Homology

[0022] In a preferred embodiment, a mutant polypeptide according to the present invention will have an amino acid sequence that is at least 50% homologous to that of a native polypeptide performing the same function as the mutant. By way of example, a KDPGal aldolase according to the present invention will have an amino acid sequence at least 50% homologous to that of any of SEQ ID NO:2, SEQ ID NO:4, and SEQ ID NO:6; in a preferred embodiment, the sequence will be at least 60% homologous thereto; in a preferred embodiment, the sequence will be at least 70% homologous thereto; in a preferred embodiment, the sequence will be at least 80% homologous thereto; in a preferred embodiment, the sequence will be at least 90% homologous thereto.

[0023] In one embodiment, a recombinant polynucleotide according to the present invention, which encodes a desired polypeptide, will be any that encodes a polypeptide having homology to a native polypeptide of the same function, as described above. In one embodiment, a recombinant polynucleotide according to the present invention, which encodes a desired polypeptide, will have an amino acid sequence that is more than 80% homologous to that of a native polynucleotide encoding a polypeptide performing the same function as the mutant. In a preferred embodiment, the polynucleotide will be at least 85% homologous thereto; in a preferred embodiment, the polynucleotide will be at least 90% homologous thereto; in a preferred embodiment, the polynucleotide will be at least 95% homologous thereto.

[0024] Sequence homology refers to the degree of identicality between two sequences of amino acid residues, or between two sequences of nucleobases. This may be determined by visual comparison of two sequences, or by use of bioinformatic algorithms that align sequences for comparison or that determine percent homology among compared sequences. Useful automated algorithms are available in the GAP, BESTFIT, FASTA, and TFASTA computer software modules of the Wisconsin Genetics Software Package (available from Genetics Computer Group, Madison, WI, USA). The alignment algorithms automated in those modules include the Needleman & Wunsch, the Pearson & Lipman, and the Smith & Waterman sequence alignment algorithms. Other useful algorithms for sequence alignment and homology determination are automated in software including: FASTP, BLAST, BLAST2, PSIBLAST, and CLUSTAL V; see, e.g., N.P. Brown et al., *Bioinformatics:* Applications Note, 1998, *14*:380-81; the U.S. National Center for Biotechnology Information at http://www.ncbi.nlm.nih.gov/Tools/index.html; and U.S. Patent No. 6,790,639, which provides software parameter settings useful for homology determination herein.

[0025] The sequence homology exhibited by nucleobase polymers, such as nucleic acids and nucleic acid analogs, may be determined by hybridization assays between a first sequence and the complement of a second sequence. Any of the well known hybridization assays may be used for this purpose, and examples of these include those described in U.S. Patent Nos. 6,767,744, and 6,783,758, with "high stringency" hybridization conditions being as defined therein.

Conservative Substitutions

[0026] In addition, conservative amino acid substitutions may be present in a polypeptide according to the present invention. The term "conservative amino acid substitution" indicates any amino acid substitution for a given amino acid residue, where the substitute residue is so chemically similar to that of the given residue that no substantial decrease in polypeptide function (e.g., enzymatic activity) results. Conservative amino acid substitutions are commonly known in the art and examples thereof are described, e.g., in U.S. Patent Nos. 6,790,639, 6,774,107, 6,194,167, or

5,350,576. In a preferred embodiment, a conservative amino acid substitution will be any one that occurs within one of the following six groups

1. Small aliphatic, substantially non-polar residues: Ala, Gly, Pro, Ser, and Thr;
2. Large aliphatic, non-polar residues: Ile, Leu, and Val; Met;
3. Polar, negatively charged residues and their amides: Asp and Glu;
4. Amides of polar, negatively charged residues: Asn and Gln; His;
5. Polar, positively charged residues: Arg and Lys; His; and
6. Large aromatic residues: Trp and Tyr; Phe.

In a preferred embodiment, a conservative amino acid substitution will be any one of the following, which are listed as Native Residue (Conservative Substitutions) pairs: Ala (Ser); Arg (Lys); Asn (Gln; His); Asp (Glu); Gln (Asn); Glu (Asp); Gly (Pro); His (Asn; Gln); Ile (Leu; Val); Leu (Ile; Val); Lys (Arg; Gln; Glu); Met (Leu; Ile); Phe (Met; Leu; Tyr); Ser (Thr); Thr (Ser); Trp (Tyr); Tyr (Trp; Phe); and Val (Ile; Leu).

[0027] Just as a polypeptide may contain conservative amino acid substitution(s), a polynucleotide hereof may contain conservative codon substitution(s). A codon substitution is considered conservative if, when expressed, it produces a conservative amino acid substitution, as described above. Degenerate codon substitution, which results in no amino acid substitution, is also useful in polynucleotides according to the present invention. Thus, e.g.,. a polynucleotide encoding a selected polypeptide useful in an embodiment of the present invention may be mutated by degerate codon substitution in order to approximate the codon usage frequency exhibited by an expression host cell to be transformed therewith, or to otherwise improve the expression thereof.

Production of Shikimate and its Precursors

[0028] In various embodiments, methods of this invention comprise the production of shikimate and shikimic acid according to the reaction schemes set forth in Figure 2. As referred to in Figure 2, "G3P" is D-glyceraldehyde 3-phosphate; "E4P" is D-erythrose 4-phosphate; "KDPGal" is 2-keto-3-deoxy-6-phosphogalactonate; "DAHP" is 3-deoxy-D-arabino-heptulosonate-7-phosphate. The enzymes used in the depicted method are (a) KDPGal aldolase (DgoA; also called 2-dehydro-3-deoxy-6-

phosphogalactonate aldolase; E.C. 4.1.2.21); (b) DAHP synthase (AroF, AroG, AroH); (c) 3-dehydroquinate synthase (AroB); (d) 3-dehydroquinate dehydratase (AroD); and (e) shikimate dehydrogenase (AroE). In other embodiments, methods of this invention comprise the production of at least one of DAHP, DHQ (3-dehydroquinate), or DHS (3-dehydroshikimate).

[0029] By catalyzing the reversible cleavage of KDPGal to pyruvate and D-glyceraldehyde 3-phosphate (G3P, Figure 2), KDPGal aldolase enables microbes (e.g., *E. coli*) to use D-galactonate as a sole carbon source. In one embodiment, *E. coli dgoA*-encoded KDPGal aldolase is overexpresssed, partially purified, and incubated with pyruvate, D-erythrose 4-phosphate (E4P), 3-dehydroquinate synthase, and 3-dehydroquinate dehydratase to form 3-dehydroshikimate. Dehydratase-catalyzed dehydration of 3-dehydroquinate provides in product 3-dehydroshikimate a chromophore suitable for continuous spectrophotometric assay.

[0030] In one embodiment, KDPGal aldolase is incubated with pyruvate and D-erythrose 4-phosphate to form DAHP; in one embodiment, KDPGal aldolase is incubated with pyruvate, D-erythrose 4-phosphate, and 3-dehydroquinate synthase to form DHQ. In one embodiment, KDPGal is expressed in a cell that provides pyruvate and E4P; DAHP can be formed thereby. In one embodiment, KDPGal is expressed in a cell that provides pyruvate, E4P,and 3-dehydroquinate synthase; DHQ can be formed thereby.

[0031] Production systems according to the present invention may be in vivo systems or in vitro systems. In vitro systems include, e.g., batch enzyme suspensions or (adsorbed or covalently) immobilized enzyme bioreactors. In vivo systems include, e.g., immobilized cell bioreactors, continuous fermentations, and batch fermentations. A DAHP synthesis system according to the present invention will include at least one KDPGal aldolase (DgoA) and a source of pyruvate and E4P. A DHQ synthesis system according to the present invention will include at least at least one KDPGal aldolase, a source of pyruvate and E4P, and at least one DHQ synthase (AroB). A DHQ synthase-containing enzymatic pathway or production system according to the present invention offers the additional benefit that, in contrast to the reversible reaction forming DAHP from pyruvate and E4P, the reaction forming DHQ from DAHP is irreversible due to cleavage of the phosphate

ester. This can result in increased yields of DHQ and downstream derivatives thereof, e.g., shikimate. The coding sequence of an exemplary *aroB* gene is the *E. coli* sequence (SEQ ID NO:7).

[0032] Recombinant host cells according to the present invention are capable of expressing at least one recombinant KDPGal aldolase and optionally at least one DHQ synthase or other shikimate pathway enzyme. In a preferred embodiment, the recombinant cell capable of expressing KDPGal aldolase, and optionally of expressing DHQ synthase, will be a walled cell. Examples of walled cells include plant cells, yeast/fungal cells, bactrerial cells, Archaea cells, and some protests. In a preferred embodiment, the recombinant cell will be a prokaryotic cell. In a preferred embodiment, the recombinant cell will be a bacterial cell. In a preferred embodiment, the recombinant cell will be a proteobacterial cell. Preferably, the recombinant host cell will lack the ability to express a functional DAHP synthase. In a preferred embodiment, the cell will be an *aroFGH* cell.

[0033] A DHS synthesis system according to the present invention will include at least at least one KDPGal aldolase, a source of pyruvate and E4P, at least one DHQ synthase, and at least one 3-dehydroquinate dehydratase. A shikimate synthesis system according to the present invention will include at least at least one KDPGal aldolase, a source of pyruvate and E4P, at least one DHQ synthase, at least one 3-dehydroquinate dehydratase, and at least one shikimate dehydrogenase.

EXAMPLES

Materials and Methods

Cloning, Plasmid Construction, Host Cells, and Transformation

[0034] Standard protocols can be used for construction, purification, and analysis of plasmid DNA. See, e.g., Sambrook, J.; Fritsch, E. F.; Maniatis, T., "Molecular Cloning – A Laboratory Manual," Cold Spring Harbor Laboratory: Plainview, NY, 1990. In various embodiments, *Escherichia coli* strain DH5α serve as the host strain for all plasmids constructions. *Klebsiella pneumoniae* genomic DNA can be purchased from the American Type Culture Collection (ATCC 700721D). *E. coli* strain CB734 [C600 *leu thi1* Δ(*gal-aroG-nadA*)50 *aroF*::*cat*(Cm^R) Δ*aroH*::Kan^R *recA1*] may

be obtained from Professor Ronald Bauerle (Univeristy of Virginia). *E. coli* strain JC7623, BW25141/pKD3 and BW25141/pKD46 may be obtained from the *E. coli* genetic stock center at Yale University. See, e.g., Lloyd, R. G.; Buckman, C., *J. Bacteriol.* 1985, *164*, 836-844; and. Datsenko, K. A.; Wanner, B. L., *Proc. Natl. Acad. Sci. USA* 2000, *97*, 6640-6645. *Taq* polymerase, large fragment of DNA polymerase I, calf intestinal alkaline phosphotase and pCR2.1-TOPO vector can be purchased from Invitrogen, *Pfu* polymerase was purchased from Strategene. DNaseI can be purchased from Roche Diagnostics. L-lactic dehydrogenase was purchased from Sigma. DNA clean and concentrator kit can be purchased from Zymo Research (Orange, CA). Phage P1 transduction, transformation with $CaCl_2$ and PCR amplifications can be performed by standard methods. *E. coli* genomic DNA can be purified as previously described. See, e.g., Pitcher, D. G.; Saunders, N. A.; Owen, R. J. *Lett. Appl. Microbiol.* 1989, *8*, 151-156. *E. coli* strains W31105, AB32486, and KL37 and cloning vectors pJF118EH8 and pTrc99A9 can be obtained from the laboratory of Dr. John Frost, Michigan State University.

[0035] *Escherichia coli dgoA* gene sequences can be obtained from the National Center for Biotechnology Information (NCBI). See, e.g., Babbitt, P. C.; Mrachko, G. T.; Hasson, M. S.; Huisman, G. W.; Kolter, R.; Ringe, D.; Petsko, G. A.; Kenyon, G. L.; Gerlt, J. A. *Science* 1995, *267*, 1159-1161. *Klebsiella pneumoniae dgoA* gene sequence can be obtained from the Genome Sequencing Center at Washington University using BLAST search against *E. coli dgoA* gene sequence.

[0036] Restriction maps for plasmids among those useful herein are set forth in Figure 3. Sites are abbreviated as follows: X=*Xba*I, B=*Bam*HI, Bg=*Bgl*II, E=*Eco*RI, P=*Pst*I, K=*Kpn*I, M=*Mfe*I, N=*Nco*I, S=*Sma*I. Parentheses indicate that the designated enzyme site has been eliminated. Lightface lines indicate vector DNA; Boldface lines indicate insert DNA.

pNR5.223:

[0037] The 0.6-kb fragment containing the *E. coli dgoA* gene and its ribosomal binding site are amplified from *E. coli* W3110 genomic DNA using *Taq* polymerase with the following pair of primers containing *Xba*I recognition sequences, JWF 430 5'-GCTCTAGATGCAGTGGCAAACTAAACT (SEQ ID NO:13) and JWF 449 3'-

GACTCTAGATCATTGCACTGCCTCTCGAT (SEQ ID NO:14). The PCR fragment is mixed with pCR2.1TOPO vector to afford the 4.5-kb plasmid pNR5.223.

pNR7.088:

[0038] The 0.6-kb fragment containing the *E. coli dgoA* gene and its ribosomal binding site was amplified from *E. coli* W3110 genomic DNA using *Pfu* polymerase with the following pair of primers, JWF 484 5'-GACGGATCCTATAAGGAGCATCGCTCATG (SEQ ID NO:15), JWF 529 3'-GAAGCTGCAGTCATTGCACTGCCTCTCGAT (SEQ ID NO:16). The PCR primers were designed to include *Bam*HI and *Pst*I recognition sequences at the 5' and 3' ends, respectively of the *dgoA* gene. Localization of the amplified *dgoA* as a *Bam*HI-*Pst*I fragment into the corresponding sites of pTrc99A afforded the 4.8-kb plasmid pNR7.088.

pNR7.118:

[0039] The 0.6-kb *E. coli dgoA* locus is excised from pNR7.088 by *Bam*HI and *Pst*I double digestion. Ligation to pJF118EH affords plasmid pNR7.118 in which the *E. coli dgoA* locus was located downstream the tac promoter.

pNR6.252:

[0040] The 5.9-kb plasmid contains *K. pneumoniae dgoA* gene located behind the tac promoter of pJF118EH. The 0.6-kb *dgoA* gene with its ribosomal binding site is amplified from *Klebsiella pneumoniae* subsp. *pneumoniae* genomic DNA (ATCC 700721D) using the following primers, JWF 501 5'-GACAGGAATAAGGAGCATCG (SEQ ID NO:17), and JWF 499 5'-GGAGGTAAACGGTACGTGGT (SEQ ID NO:18). The resulting PCR fragment is ligated into pCR2.1TOPO vector by TA TOPO cloning technique to afford pNR6.223B. The 0.6-kb locus is then excised from pNR6.223B by *Eco*RI digestion and ligated to the *Eco*RI site of pJF118EH to afford plasmid pNR6.252.

pNR7.126:

[0041] The 6.4-kb plasmid is constructed by ligating the *aroF*^{FBR} ("*aroF* Feedback-Resistant") gene with its ribosomal binding site into the *Eco*RI site of pJF118EH. The *aroF*^{FBR} fragment is amplified by PCR from pKD12.112 using the following primers containing *Eco*RI terminal recognition

sequence, JWF541 5'-GGAATTCGCATAAACAGGATCGCCATCA (SEQ ID NO:19) and JWF542 5'-CTGGATCCTTAAGCCACGCGAGCCGT (SEQ ID NO:20). See Draths, K.M. et al., *J. Am. Chem. Soc.*, 1999, *121*, 1603-04.

pNR7.288:

[0042] The 0.8-kb *cat* gene is excised from pSU18 by digestion with *Bam*HI, then is inserted into the *Bgl*II site internal to the *aroF*^{FBR} gene in pKD12.112 to produce plasmid pNR7.288.

pNR7.297:

[0043] The 1.3-kb fragment containing *aroG* gene is amplified from *E. coli* W3110 genomic DNA using the following primers containing *Bam*HI terminal recognition sequence, JWF 610 5'-GTGGATCCTTAATCCGTTCATAGTGTAAA (SEQ ID NO:21), and JWF 611, 5'-TGGGATCCATGAGAAAGCCGACTGCAA (SEQ ID NO:22). The PCR fragment is ligated into pSU18 vector to create pNR7.260. Ligation of a *Ssp*I/*Ava*I digested *tet*-encoding fragment of DNA obtained from plasmid vector pBR322 into the *Mfe*I site internal to the *aroG* gene in pNR7.260 resulted in plasmid pNR7.297.

pNR7.290:

[0044] The 1.3-kb fragment containing *aroH* gene is amplified from *E. coli* W3110 genomic DNA using the following primers JWF-625, 5'-GTTCGTCAGTGCAGGATGGA (SEQ ID NO:23) and JWF-626, 5'-GTTCAGGCGTGAGTTTTCTGCT (SEQ ID NO:24). The PCR product is initially cloned into plasmid vector pCR2.1-TOPO, then the fragment containing *aroH* is digested with *Hin*dIII/*Xba*I and inserted into the *Hin*dIII/*Xba*I site of plasmid pTrc99A to afford a 5.5-kb plasmid pNR7.289B. A 1.3-kb *Kan*^{R} gene excised from plasmid pKAD62A by digestion of *Pst*I is cloned into the *Pst*I site internal to the *aroH* gene in pNR7.289B to yield plasmid pNR7.290.

pNR8.074:

[0045] The 1.6-kb *serA* gene is excised from plasmid pRC1.55B by *Sma*I digestion. Ligation of *serA* gene to plasmid pNR6.252 predigested with *Hin*dIII to afford the 7.5-kb plasmid pNR8.074.

pNR8.075:

[0046] The 1.6-kb *serA* gene is excised from plasmid pRC1.55B by *Sma*I digestion. Ligation of *serA* gene to plasmid pNR7.088 predigested with *Hin*dIII afforded the 6.8-kb plasmid pNR8.075.

Construction of *aroFaroGaroH* Triple Mutant Strain NR7:

[0047] Disruption of *aroF*, *aroG* and *aroH* genes in *E. coli* KL3 is done as follows. Plasmid pNR7.288 is digested with *Eco*RI to liberate the insertionally inactivated *E. coli aroF* gene. The purified fragment (*aroF*::*cat*) is electroporated into the hyper-recombinant *E. coli* strain JC7623. Chloroamphenicol resistant colonies are picked up from LB plate containing 20 μg/mL chloroamphenicol, and the genotype verified by size analysis of the DNA fragment amplified by PCR from chromosomal DNA using the following primers JWF-541, 5'-GGAATTCGCATAAACAGGATCGCCATCA (SEQ ID NO:25), and JWF-542, 5'-CTGGATCCTTAAGCCACGCGAGCCGT (SEQ ID NO:26). The *aroF*::*cat* mutation is then transferred from JC7623*aroF*::*cat* to *E. coli* KL3 by phage P1-mediated transduction to prepare KL3*aroF*::*cat*.

[0048] Similarly, the fragment comprising *aroH*::*Kan*^{R} is excised from plasmid pNR7.290 by digestion with *Xba*I and *Hin*dIII, then electroporated into JC7623. The JC7623*aroH*::*Kan*^{R} mutants resulting from homologous recombination are resistant to 50 μg/mL kanamycin. The correct genotype is verified by PCR analysis from chromosomal DNA using the following primers JWF-636, 5'-TCCGTACTGCGCGTATTGAGA (SEQ ID NO:27) and JWF-637, 5'-AGAGGCGAGTTTTTCGACCA (SEQ ID NO:28). P1 phage mediated transduction employed to transfer the *aroH*::*Kan*^{R} mutation into KL3*aroF*::*cat* and produce NR3 [KL3*aroF*::*cat aroH*::*Kan*^{R}].

[0049] The *aroG* mutation is generated by the methods described by Datsenko and Wanner. See, Datsenko, K. A.; Wanner, B. L. *Proc. Natl. Acad. Sci. USA* 2000, *97*, 6640-6645. Plasmid pNR7.297 is digested with *Kpn*I and *Pst*I to liberate a fragment containing *aroG*::*tet* cassette, the purified DNA is electrporated into NR3/pKD46 while expressing Red recombinase. Recombinants are selected for tetracycline resistance (5 μg/mL). Plasmid pKD46 are eliminated by growth at 42°C. Disruption of *aroG* is confirmed by PCR analysis using the following primers JWF-669, 5'-GCAGCATTGTGCCGCCAGAA (SEQ ID NO:29) and JWF-670, 5'-

GTGCGCTGGTGAAATATCTT (SEQ ID NO:30). The KL3*aroF*::*cat aroG*::*tet aroH*::*Kan*^{R} strain is designated herein as *E. coli* NR7.

Enzyme Assay

[0050] The cells are suspended in KH_2PO_4 (20 mM, pH 6.5) containing PMSF (1 Mm; phenyl methyl sulfonyl fluoride). Cell lysis is achieved by two passes through a French pressure cell (SLC Aminco) at 16000 psi (110.3 MPa). Cell debris is separated from lysate by centrifugation at 48000g for 20 min at 4°C. Protein concentration is quantified using the Bradford dye-binding procedure with the assay solution purchased from Bio-Rad. See Bradford, M.M., *Anal. Biochem.*, 1976, *72*, 248-54.

KDPGal Cleavage Assay

[0051] KDPGal aldolase activity is determined using a coupled enzyme assay previously described by Meloche: Meloche, H.P., Wood, W.A., *J. Biol. Chem.*, 1964, *239*, 3515-18.. 2-keto-3-deoxy-6-phosphogalactonate (KDPGal) was prepared following the method of Toone: Toone, E.J. et al., *J. Mol. Catal. B-Enzymatic*, 1998, *5*, 103-11. To a 1 mL quartz cuvette were added 954 µL KH_2PO_4 buffer (20 mM, pH 7.5), 10 µL NADH (35 mM), 10 µL L-lactic dehydrogenase (L-LDH, EC 1.1.1.27, type II, rabbit muscle, 1U/µL) and 10 µL appropriately diluted cellular lysate (1:10 to 1:50 dilution in 20 mM KH_2PO_4, pH 7.5). The solution is mixed and pre-incubated for 2 min at room temperature. The reaction is initiated by addition of KDPGal (16 µL, 100 mg/mL, Li^+ salt). The absorbance at 340 nm is recorded continuously for 1 min. One unit of KDPGal aldolase activity is defined by the catalyzed loss of one µmol of NADH per minute.

DAHP Formation Assay

[0052] Enzyme activity is measured by following the formation of 3-dehydroshikimic acid (DHS) over time when DHQ synthase (3-dehydroquinate synthase, E.C. 4.2.3.4) and DHQ dehydratase (E.C. 4.2.1.10) are utilized as coupling enzymes. DHQ synthase is prepared as described in Frost, J.W. et al., *Biochemistry*, 1984, *23*, 4470-75; DHQ dehydratase is purified from *E. coli* AB2848/pKD201 The reaction (1 mL) contains 50 mM morpholinepropanesulfonic acid (MOPS) buffer (pH 7.5), 1 mM D-erythrose 4-phosphate, 1 mM pyruvate, 50 µM $CoCl_2$, 10 µM NAD, 1U of DHQ synthase, 1U of DHQ dehydratase and cellular lysate. The reaction is initiated by

addition of diluted cellular lysate to the assay solution, and the absorbance at 234 nm monitored continuously for 5 min. One unit of evolved enzyme activity is defined by the formation of one µmol of DHS per minute.

Error-Prone PCR

[0053] Random mutagenesis of *dgoA* gene is conducted using methods described by Cadwell and Joyce: Cadwell, R.C., Joyce, G.F., *PCR Meth. Appl.*, 1992, *2*, 28-33. PCR is performed in a 100 µL reaction mixture containing 10 ng of *dgoA* fragment as template, 20 mM Tris-HCl, pH 8.4, 50 mM KCl, 7 mM $MgCl_2$, 0.5 mM $MnCl_2$, 0.2 mM dATP, 0.2 mM dGTP, 1.0 mM dCTP, 1.0 mM dTTP, 50 pmol of each primer and 5U *Taq* polyerase (Invitrogen). Conditions for PCR are as follows: one cycle of 4 min at 94°C, 22 cycles of 45 sec 94°C, 45 sec 45°c, 45 sec 72°C; and one cycle of 10 min 25°C.

DNA Shuffling

[0054] DNA shuffling is performed following the protocol of Stemmer modified by Zhao and Arnold. See, Stemmer, W. P. C. *Proc. Natl. Acad. Sci. USA* 1994, *91*, 10747-10751; Stemmer, W. P. C. *Nature* 1994, *370*, 389-391; and Zhao, H.; Arnold, F. H. *Nucleic Acids Res.* 1997, *25*, 1307-1308. The 0.6-kb *dgoA* gene of interest is amplified using *Pfu* polymerase under standard PCR conditions and cleaned through DNA Clean and Concentrator kit. Fragments for shuffling are created by digesting the cleaned PCR product with DNaseI in a 50 µL reaction containing 5 µg DNA, 50 mM Tris-HCl, pH 8.0, 10 mM $MnCl_2$, and 0.05U of DNaseI for 10 min at 15°C. The reaction is stopped with addition of 15 µL EDTA (100 mM, pH 8.0) and 12 µL Endostop. Fragments of 20 to 80 bps are purified from 2.0% low melting point agarose gel (Invitrogen) using DE81 ion-exchange paper (Whatman). The purified DNA fragment are dissolved into 30 µL sterile water.

[0055] The fragments are reassembled by PCR without primers in a 50 µL reaction containing 20 mM Tris-HCl, pH 8.4, 50 mM KCl, 2 mM $MgCl_2$, 200 µM each dNTPs, 10 µL of DE81 purified DNA fragments and 2.5U of *Taq* polymerase. PCR is conducted as follows: 1 min 94°C followed by 45 cycles of 30 sec 94°C, 30 sec 50°C, 30 sec 72°C, followed by 5 min 72°C and 5 min 25°C.

[0056] The 0.6-kb *dgoA* genes are reassembled by PCR with forward and reverse primers in a 100 µL reaction containing 5 µL of reassembled DNA, 20 mM Tris-HCl, pH 8.4, 50 mM KCl, 2 mM $MgCl_2$, 200 µM each dNTPs, 50 pmol primers and 5U of *Taq* polymerase.

Selection Medium for Directed Evolution of KDPGal Aldolases

[0057] Selection medium comprises Na_2HPO_4 (6 g/L), KH_2PO_4 (3 g/L), NH_4Cl (1g/L), NaCl (0.5 g/L), glucose (4 g/L), $MgSO_4$ (0.12 g/L), thiamine (6 mg/L), L-leucine (25 mg/L), nicotinic acid (6 mg/L). IPTG is added to a final concentration of 0.2 mM or 0.05 mM. L-Phenylalanine (40 mg/L), L-tyrosine (40 mg/L) and L-tryptophan (40 mg/L) are added as indicated. Solid medium is prepared by addition of 1.5% (w/v) Difco agar to medium solution.

Specific Example 1

Purification of KDPGal Aldolase from *E. coli* and In-Vitro Biosynthesis of 3-Dehydroshikimate

[0058] *E. coli* AB3248/pNR5.223 is grown in LB medium containing 50 µg/mL ampicillin at 37°C. IPTG (isopropyl-beta-D-thiogalactopyranoside) is added to a final concentration of 0.2 mM when OD600 reached 0.5. The cells are grown for an additional 6 h, and pelleted by centrifugation (4200g, 10 min). The cells are washed with 0.9% NaCl solution once and suspended in 20 mM KH_2PO_4, pH 6.5 with 1 mM PMSF. Disruption of the cells is achieved by two passages through a French pressure cell (16000 psi; 110.3 MPa). Cell debris is removed by centrifugation at 48000g for 20 min. The resulting crude cell-free lysate is first treated with protamine sulfate and solid $(NH_4)_2SO_4$, then applied to a DEAE-cellulose (Whatman) column. The column is washed with a gradient mixture of buffer B (20 mM KH_2PO_4, 50 mM KCl, pH 7.5) and buffer C (20 mM KH_2PO_4, 400 mM KCl, pH 7.5). Fractions containing KDPGal aldolase are concentrated, dialyzed, quick frozen in liquid nitrogen and stored at –80°C (87 units/mg).

[0059] D-Erythrose 4-phosphate (0.45 mL, 12 mM, pH 7.0), sodium pyruvate (0.054 mL, 100 mM, pH 7.0), $CoCl_2$ (0.027 mL, 10 mM), NAD (0.054 mL, 1 mM), DHQ synthase (2 units) and DHQ dehydratase (2 units) are incubated with KDPGal aldolase (100 units) from *E. coli* at room

temperature for 2 h. Protein is subsequently removed by ultrafiltration using Millipore PM-10 membrane. 3-Dehydroshikimate was formed in 90% yield, as determined by ^{1}H NMR analysis.

Specific Example 2

Cloning and Characterization of *dgoA* Genes from Other Bacterial Sources.

[0060] *E. coli* wild-type *dgoA*-encoded KDPGal aldoalse showed weak activity toward accepting D-erythrose 4-phosphate as a substrate. KDPGal aldolases from other bacterial sources might have higher activities for catalyzing the condensation of pyruvate with D-erythrose 4-phosphate. Obtaining *dgoA* genes from other bacteria would also enable a cross-species DNA family shuffling that has been reported to improve enzyme performance rapidly. Although KDPGal aldolase activities have been identified in *Pseudomonas saccarophila, Pseudomonas cepacia, Caulobacter cresentus, Azotobacter vinelandii, Rhizobium meliloti, Gluconobacter liquefaciens,* and nonpathogenic *Mycobacteria*, none of these *dgoA* gene sequences was known except in *Caulobacter cresentus* in which the genomic sequence has been obtained. See: Crameri, A. et al., *Nature* 1998, *391*, 288-291; Kurn, N. et al., *J. Bacteriol.* 1978, *135*, 517-520; Wong, T.Y.; Yao, X., *Appl. Environ. Microbiol.* 1994, *60*, 2065-2068; Arias, A.; Cervenansky, C., *J. Bacteriol.* 1986, *167*, 1092-1094; Stouthammer, A. H., *Biochim. Biophys. Acta* 1961, *48*, 484-500; Szumilo, T., *Mycobacteria. J. Bacteriol.* 1981, *148*, 368-370; and Nierman, W.C. et al., *Proc. Natl. Acad. Sci. USA* 2001, *98*, 4136-4141.

[0061] Performing a BLAST (Basic Local Alignment Search Tool) search against *E. coli dgoA* nucleotides sequence in microbial genome database only yielded two possible *dgoA* sequences from *Klebsiella pneumoniae* and *Salmonella typhimurium* LT2. BLAST search against the *E. coli dgoA* protein sequence afforded several more hits including *Caulobacter cresentus* CB15, *Agrobacterium tumefaciens, Ralstonia solanacearum, Bradyrhizobium japonicum, Brucella melitensis* and *Sinorhizobium meliloti.* The genomic DNA of *K. pneumoniae, S. typhimurium* LT2 (ATCC 15277), *A. tumefaciens* (ATCC 17805) and *C. cresentus* CB15 (ATCC 19089) are readily available from the American Type Culture Collection (ATCC). Thus, the open reading frames of the *K. pneumoniae, S. typhimurium* LT2, *A. tumefaciens*

and *C. cresentus* CB15 *dgoA* genes were amplified from their respective genomic DNA using PCR and cloned into a medium copy number expression vector pJF118EH with transcription under the control of a P_{tac} promoter to prepare plasmid pNR6.252, pNR7.120, pNR6.300 and pNR7.063, respectively. The native start codon of GTG in *C. cresentus dgoA* was changed into an ATG start codon in plasmid pNR7.063. The corresponding plasmids were transformed into *E. coli* CB734, and in all cases, the KDPGal aldolase activities in crude cell lysates were confirmed and determined (Table 1).

Table 1. KDPGal Aldolases from Various Microorganisms.

dgoA source	*dgoA* size (nt)	Identity[a] with *E. coli dgoA*	KDPGal cleavage [b]	DAHP formation [b]
Escherichia coli[c]	618	100%	7.6	0.068
Klebsiella pneumoniae[d]	618	82%	77	0.29
Salmonella typhimurium[e]	618	81%	10	0.080
Agrobacterium tumefaciens[f]	630	54%	4.8	0.30
Caulobacter cresentus[g]	582	60%	3.6	0.23

[a] Identity is calculated based on nucleotide sequence using the global sequence alignment provided by Biology Workbench (http://workbench.sdsc.edu); see Pearson, W. R.; Lipman, D. J., Improved tools for biological sequence comparison, *Proc. Natl. Acad. Sci. USA* 1988, *85*, 2444-2448, and Pearson, W. R., Rapid and sensitive sequence comparison with FASTP and FASTA, *Methods Enzymol.* 1990, *83*, 63-98.

[b] Specific activity is defined as units of enzyme activity per mg of protein in crude cell lysate. One unit of activity = one μmol of KDPGal cleaved or DAHP formed per minute. [c] *E. coli* CB734/pNR7.088.

[d] *E. coli* CB734/pNR6.252.

[e] *E. coli* CB734/pNR7.120.

[f] *E. coli* CB734/pNR6.300.

[g] *E. coli* CB734/pNR7.063.

[0062] All of the bacteria identified in the BLAST searches are members of the proteobacteria. Thus, in a preferred embodiment of a recombinant KDPGal aldolase according to the present invention, the amino acid sequence of the recombinant will be a variant of the amino acid sequence of a native (i.e. wild-type) KDPGal aldolase obtained from a member of the bacteria; in a preferred embodiment, it will be obtained from a member of the proteobacteria. In methods and pathways according to the present invention, the KDPGal aldolase(s) used therein may have either a native or a mutant KDPGal aldolase amino acid sequence. Where a native KDPGal aldolase amino acid sequence is utilized, in a preferred embodiment, it will be obtained from a member of the bacteria; in a preferred embodiment, it will be obtained from a member of the proteobacteria. In a preferred embodiment, a bacterium providing such a native amino acid sequence will be a member of any one of the genera *Escherichia*, *Klebsiella*, *Salmonella*, *Caulobacter*, *Agrobacterium*, *Ralstonia*, *Bradyrhizobium*, *Brucella*, and *Sinorhizobium*.

[0063] The KDPGal aldolases of *E. coli* (SEQ ID NO:2), *K. pneumoniae* (SEQ ID NO:4), and *S. typhimurium* (SEQ ID NO:6), were all found to be 205 residues in length. The remaining KDPGal aldolases identified from BLAST analyses of genomic DNA, were found to range from 194 to 213 residues in length. In a preferred embodiment, a KDPGal aldolase according to the present invention will be about 190 to about 215 residues; in a preferred embodiment, it will be about 200 to about 210 residues; and in a preferred embodiment, it will be about 205 residues in length.

[0064] Among the five KDPGal aldolases, *K. pneumoniae* and *A. tumefaciens* KDPGal aldolases showed highest activities toward DAHP formation (see Table 1). The *dgoA* gene coding sequences of *K. pneumoniae* (SEQ ID NO:3) and *S. typhimurium* LT2 (SEQ ID NO:5) were found to have the highest nucleotide sequence homology of about 81% with that of the *E. coli dgoA* (SEQ ID NO:1) (see Table 1).

Specific Example 3

Directed Evolution of KDPGal Aldolases

[0065] The *dgoA* genes of *E. coli*, *K. pneumoniae*, and *S. typhimurium* were each evolved by use of error-prone PCR and DNA shuffling. *E. coli dgoA* and *K. pneumoniae dgoA* and were subjected to two rounds of error-prone PCR mutagenesis followed by one round of DNA shuffling. Identification of increased activity relative to condensation of pyruvate with D-erythrose 4-phosphate was based on a selection that increased in stringency with each round of directed evolution.

[0066] The activity of the resulting KDPGal aldolases in catalyzing condensation of pyruvate and D-erythrose 4-phosphate was investigated in *E. coli* CB734, which lacks all isozymes of DAHP synthase. Thus, growth of *E. coli* CB734 on glucose-containing minimal salts medium required supplementation with L-phenylalanine, L-tyrosine, L-tryptophan and aromatic vitamins (entry 1, Tables 2-4, below).

Table 2

Directed Evolution of *E. coli* KDPGal aldolase.

Entry	Construct[a]	M9[b]	M9[c]	F[d]	YF[d]	YFW[d]	YFWvit[d]
1	*E. coli* CB734	-[e]	-	-	-	-	+[f]
2	CB734/pNR7.088	-	-	-	-	+	+
3	CB734/pEC01	-	-	-	+	+	+
4	CB734/pEC02	-	+	+	+	+	+
5	CB734/pEC03	+	+	+	+	+	+

[a] All native and evolved *dgoA* genes were inserted into the same plasmid (pTrc99A) with transcription controlled by a *Ptrc* promoter. [b] Contained L-leucine and 0.05 mM IPTG. [c] Contained L-leucine and 0.2 mM IPTG. [d] Supplements added to M9 medium containing L-leucine and 0.2 mM IPTG included: F, L-phenylalanine; Y, L-tyrosine; W, L-tryptophan; vit, *p*-aminobenzoate, *p*-hydroxybenzoate, 2,3-dihydroxybenzoate. [e] no growth (-). [f] growth (+).

Table 3

Directed Evolution of *K. pneumoniae* KDPGal Aldolase

Entry	Construct[e]	M9[a]	M9[b]	F[c]	YF[c]	YFW[c]	YFWvit[c]

1	*E. coli* CB734	-[d]	-	-	-	-	+[e]
2	CB734/pNR7.118	-	-	-	-	+	+
3	CB734/pNR6.252	-	-	-	+	+	+
4	CB734/pKP01	-	-	+	+	+	+
5	CB734/pKP02	-	+	+	+	+	+
6	CB734/pKP03	+	+	+	+	+	+

[a] contained 0.05 mM IPTG. [b] contained 0.2 mM IPTG. [c] Supplements added to M9 medium containing 0.2 mM IPTG included: F, L-phenylalanine; Y, L-tyrosine; W, L-tryptophan; vit, *p*-aminobenzoate, *p*-hydroxybenzoate, 2,3-dihydroxybenzoate. [d] growth (+). [e] no growth (-). [e] All native and evolved *dgoA* genes were inserted into the same plasmid (pJF118EH) with transcription controlled by a *Ptac* promoter.

Table 4

Directed Evolution of *S. typhimurium* KDPGal Aldolase

Entry	Construct[a]	M9[b]	M9[c]	F[d]	YF[d]	YFW[d]	YFWvit[d]
1	*E. coli* CB734	-[e]	-	-	-	-	+[f]
2	CB734/pNR7.120	-	-	-	-	+	+
3	CB734/pST01	-	-	-	+	+	+
4	CB734/pST02	-	-	+	+	+	+
5	CB734/pST03	-	+	+	+	+	+
6	CB734/pST04	+	+	+	+	+	+

[a] All native and evolved *dgoA* genes were inserted into the same plasmid (pJF118EH) with transcription controlled by a P_{tac} promoter. [b] Contained L-leucine and 0.05 mM IPTG. [c] Contained L-leucine and 0.2 mM IPTG. [d] Supplements added to M9 medium containing L-leucine and 0.2 mM IPTG included: F, L-phenylalanine; Y, L-tyrosine; W, L-tryptophan; vit, *p*-aminobenzoate, *p*-hydroxybenzoate, 2,3-dihydroxybenzoate. [e] no growth (-). [f] growth (+).

[0067] It was also found that *E. coli* CB734/pNR7.118 with its plasmid-encoded *E. coli* DgoA was able to biosynthesize its own aromatic vitamins (entry 2, Table 3). Plasmid-encoded *K. pneumoniae* DgoA afforded a 4-fold higher KDPGal aldolase specific activity in *E. coli* CB734/pNR6.252

relative to plasmid-encoded *E. coli* DgoA in *E. coli* CB734/pN7.118. *E. coli* CB734/pNR6.252 was able to provide for its own aromatic vitamin and L-tryptophan requirements (entry 3, Table 3).

A. Evolution of KDPGal aldolase from *E. coli*

[0068] The *dgoA* gene encoding the native *E. coli* KDPGal aldolase is amplified under mutagenic (error-prone) PCR condition with the following primers, JWF 484 5'-GACGGATCCTATAAGGAGCATCGCTCATG (SEQ ID NO:33), JWF 529 3'-GAAGCTGCAGTCATTGCACTGCCTCTCGAT (SEQ ID NO:34). The 0.6-kb *E. coli dgoA* amplification product is purified, digested with restriction enzymes *Bam*HI and *Pst*I. The fragments are cloned into the corresponding sites of pTrc99A and transformed into competent CB734 cells by electroporation to generate the first generation plasmid library. After transformation, a library of 1×10^6 colonies are spread on minimal salts plates supplement with tyrosine and phenylalanine with 0.2 mM IPTG. After incubation at 37°C for 3 days, 50 large colonies were picked up and a plasmid mixture (pEC01-mix) are prepared. For the second round of selection, the mutant 0.6-kb *dgoA* gene is isolated and further amplified under mutagenic PCR condition, the fragments are cloned into pTrc99A vector. A library of 1×10^6 colonies is spread on minimal salts plates with 0.2 mM IPTG without any aromatic amino acids supplementation. More than 100 colonies are produced after 4 days of incubation, 50 large colonies replicated on a LB/Ap plate. After overnight incubation, the cells are scraped from the plate with sterile water and a plasmids mixture was prepared (pEC02-mix). For the third round, the 0.6-kb fragments encoding the mutant DgoA are PCR amplified from the $pEC0_2$-mix, the fragments aer shuffled to combine the beneficial point mutations and cloned into pTrc99A as above. A library of 3×10^5 colonies are spread on minimal salts plates with 0.05 mM IPTG. After 3 days at 37°C, 7 largest colonies are picked up for characterization. See Table 5, below.

Table 5

Mutations and Specific Activities of *E. coli* KDPGal Aldolase Variants

DgoA formation	Mutations	DAHP sp. activity[a] (U/mg)
Wild-type		0.086
EC03-1	F33I, D58N, Q72H, A75V, V85A, V154F	0.56
EC03-2	D30G, F33I, D34G, S42T, A75T, V85A, V154F, L179I, A182P	0.30
EC03-3	F33I, D34G, K59R, V85A, A111P, G134S, P135L, V154F, P159A	0.56
EC03-4	F33I, D34G, S42T, D74N, V85A, A122V, V154F, D167E, A190T	0.32
EC03-5	S42T, K59M, V85A, A122V, V154F, D178V	0.37
EC03-6	S42T, V85A, H90Y, V154F, L175I	0.32
EC03-7	K6N, T17M, V85A, I89T, V154F, S185P	0.29

[a] *E. coli* CB734 was used as host strain for expression of the evolved enzymes.

B. Evolution of KDPGal Aldolase from *K. pneumoniae*

[0069] *K. pneumoniae dgoA* is amplified under mutagenic PCR conditions with the following primers, JWF559 5'-GGAATTCGACAGGAATAAGGAGCATCG (SEQ ID NO:31) and JWF560 5'-GACGGATCCTCATTTCACTGCCTCTCGAT (SEQ ID NO:32). The 0.6-kb amplification product is purified through DNA Clean and Concentrator kit, followed by double digestion with restriction enzymes *Eco*RI and *Bam*HI, and cloned into the *Eco*RI-*Bam*HI restriction site of expression vector pJF118EH to generate the first generation plasmid library. The plasmid library is electroporated into CB734 and a library of 1×10^6 colonies are plated out on minimal salts plates with phenylalanine and 0.2 mM IPTG supplementation.

L-Leucine is added throughout the selection since CB734 is also a leucine-auxotrophy strain. A single colony is produced after 48 h incubation at 37°C, and the plasmid carrying the first generation mutant KP01-1 prepared from the colony. For the second generation, KP01-1 is amplified under mutagenic PCR conditions and cloned as above. A library of 6×10^5 colonies are spread on minimal salts plates with 0.2 mM IPTG and no aromatic amino acids supplementation. After 4 days incubation at 37°C, 50 largest colonies were picked up and a mixture of the plasmid was prepared. For the third round, the mutated *dgoA* genes KP02 were fragmented and shuffled to combine beneficial point mutations and subsequent cloned and transformed as before. A library of 3×10^4 colonies are spread on minimal salts plates with 0.05 mM IPTG. After 3 days at 37°C, 7 largest colonies were picked up for characterization. See Table 6 below.

Table 5

Mutations and Specific Activities of *K. pneumoniae* KDPGal Aldolase Variants

DgoA	Mutations	DAHP sp. activity[a] (U/mg)
Wild-type		0.29
KP03-1	I10V, V85A, V154F, E187D, F196I	0.80
KP03-2	I10V, P70L, V85A, P106S, V154F, S185L, F196I	0.15
KP03-3	I10V, E71G, V85A, P106S, V154F, E187D, Q191H, F196I	1.30
KP03-4	I10V, V85A, V154F, A195T, F196I	0.51
KP03-5	I10V, I16V, P70L, V85A, R96Q, P106S, V154F, F196I	0.049
KP03-6	I10V, V85A, V154F, F196I	0.66
KP03-7	I10V, V85A, P106S, V154F, F196I	0.65

[a] *E. coli* CB734 was used as host strain for expression of the evolved enzymes.

C. Evolution of KDPGal Aldolase from *S. typhimurium*.

[0070] *S. typhimurium dgoA* was subjected two rounds of error-prone PCR mutagenesis and two rounds of DNA shuffling. The plasmid library from the first round of error-prone PCR mutagenesis was

electroporated into *E. coli* CB734 and plated out onto minimal salts plates containing L-tyrosine, L-phenylalanine and 0.2 mM IPTG. *E. coli* CB734/pST01 (entry 3, Table 4) colonies resulted from the first round of PCR mutagenesis performed using wild-type *S. typhimurium dgoA* as template only required L-tyrosine and L-phenylalanine supplementation for growth. The second round of PCR mutagenesis gave *E. coli* CB734/pST02 colonies (entry 4, Table 4) whose growth required only L-phenylalanine supplementation. The third round of mutagenesis involving shuffling gave *E. coli* CB734/pST03 colonies that grew in the absence of aromatic amino acids supplements (entry 5, Table 4). The fourth round of mutagenesis involving shuffling gave CB734/pST04 colonies that grew in minimal salts medium without aromatic amino acids supplementation at reduced KDPGal aldolase expression level by lowering IPTG concentration (entry 6, Table 4). The *dgoA* gene variants from seven largest colonies after the final round of selection were sequenced and their encoding KDPGal aldolase activities toward DAHP formation were characterized (Table 7). All seven evolved KDPGal aldolase showed higher activity toward DAHP formation activity as compared to the wild-type *S. typhimurium* KDPGal aldolase. The most active mutant ST04-5 showed a 15-fold increase in activity.

Table 7

Mutations and Specific Activities of *S. typhimuriume* KDPGal Aldolase Variants.

DgoA	Mutations	DAHP activity[a] (U[b]/mg)
Wild-type		0.080
ST04-1	V28L, S42T, S50P, P150L, L175S	0.48
ST04-2	V28M, S42T, S50P, P150L, D178G, N198K	0.85
ST04-3	D20E, V28L, S42T, L175S	0.84
ST04-4	V28M, S42T, Q123R, T158M, N161D, D178G	0.54
ST04-5	D20E, V28M, S42T, I89T, P150L, D178G	1.24
ST04-6	V28M, S42T, S50P, Q164A, L175S, N198K	0.42
ST04-7	V28L, S42T, P91Q, P150L, T158M, D178G,	1.04

N198K

[a] *E. coli* CB734 was used as host strain for expression of evolved enzymes. [b] One unit of DAHP synthase catalyzes the formation of one μmol of 3-dehydroshikimate per minute at 25°C.

D. Characterization of *E. coli*, *K. pneumoniae* and *S. typhimurium* DgoA Mutants

[0071] After directed evolution, a total of twenty-one active mutants selected from the *E. coli*, *K. pneumoniae* and *S. typhimurium* DgoA mutants were further characterized. Each mutant contained 4-9 amino acids substitutions. No insertion or deletion mutants were found. Two amino acids substitutions (V85A, V154F) were observed in all of the seven most active *K. pneumoniae dgoA* and seven most active *E. coli dgoA* mutants. However, these two mutations were not found in any of the seven most active *S. typhimurium* mutants. Instead, all seven of the most active *S. typhimurium* mutants contained a S42T substitution, as did four of the seven most active *E. coli* mutants. EC03-1, the most active evolved *E. coli* KDPGal aldolase, exhibited an 8-fold higher DAHP formation specific activity and a 7-fold reduced KDPGal cleavage specific activity relative to the native *E. coli* KDPGal aldolase (entry 2, Table 8). KP03-3, the most active evolved *K. pneumoniae* KDPGal aldolase, showed a 4-fold higher DAHP formation specific activity and a 30-fold reduced KDPGal cleavage specific activity relative to native *K. pneumoniae* KDPGal aldolase (entry 4, Table 8). ST04-5, the most active evolved *S. typhimurium* KDPGal aldolase, exhibited a 15-fold higher DAHP formation specific activity and a 2-fold reduced KDPGal cleavage specific activity relative to wild-type *S. typhimurium* KDPGal aldolase (entry 6, Table 8).

Table 8

Specific Activities of Wild-Type and Evolved KDPGal Aldolase Isozymes.

Enzyme	DAHP assay[a] (U/mg)	KDPGal assay[a] (U/mg)
E. coli DgoA[b]	0.068	6.7
EC03-1[c]	0.56	1.0

K. pneumoniae DgoA[d]	0.29	77
KP03-3[e]	1.30	2.6
S. typhimurium DgoA[f]	0.080	11
ST04-5[g]	1.24	4.8

[a] Specific activity is defined as units of enzyme activity per mg of protein in crude cell lysates. One unit of activity = one μmol of KDPGal cleaved or DAHP formed per minute. Crude cell lysates were prepared from [b] *E. coli* CB734/pNR7.088; [c] *E. coli* CB734/pEC03-1; [d] *E. coli* CB734/pNR6.252; [e] *E. coli* CB734/pKP03-3; [f] *E. coli* CB734/pNR7.120; [g] *E. coli* CB734/pST04-5.

[0072] In addition to V85A and V154F, two other amino acid substitutions (I10V, F196I) were found in all seven of the most active *K. pneumoniae* mutants. Only one substitution was found in all seven of the most active *S. typhimurium* mutants; however, a conservative substitution at V28 (either V28M or V28L) was also found in all seven mutants. One amino acid substitution (P70L) was observed solely in the two active *K. pneumoniae* mutants identified as exhibiting less activity than wild-type *K. pneumoniae* DgoA; however, these two mutants still exhibited significant aldolase activity, with one of them (KP03-2) exhibiting greater DAHP formation activity than either of the *E. coli* or *S. typhimurium* wild type enzymes.

[0073] In sum, the following amino acid mutations were identified as being associated with improved mutants of the *E. coli*, *K. pneumoniae*, and *S. typhimurium* DgoA enzymes: X10V, X28L or X28M, X42T, X85A, X154F, and X196I ("X" representing any amino acid residue at that position), more specifically, I10V, V28L or V28M, S42T, V85A, V154F, and F196I. Thus, in a preferred embodiment of a recombinant KDPGal aldolase according to the present invention, the KDPGal aldolase will contain at least one of said mutations. Likewise, the following amino acid mutation was identified as being associated with impaired, though active, mutants of the *K. pneumoniae* enzyme: X70L, more specifically, P70L. Thus, in a preferred embodiment of a recombinant KDPGal aldolase according to the present invention, the KDPGal aldolase will contain no mutation that is X70L.

[0074] The mutations were characterized at the DNA level as follows, with nucleotide substitutions and codon substitutions shown in parentheses:

I10V (a28g; atc28-30gtc); V28M (g82a; gtg82-84atg); V28L (g82t; gtg82-84ttg); S42T (t124a; tcc124-126acc); V85A (t254c; gtt253-255gct); V154F (g460t; gtt460-462ttt); and F196I (t586a; ttc586-588atc).

Specific Example 5

DgoA Family Shuffling

[0075] Family shuffling of all five *dgoA* genes constitutes one option for improving DAHP formation activity. However, a major limitation cited for family shuffling of homologous genes is its reliance on PCR-based assembly of short random fragments generated from homologous genes. This demands a level of sequence identity of more than 70% and 10-15 bp stretches of continuous sequence identity between sequences in order for recombination to occur. Therefore, only *K. pneumoniae* and *S. typhimurium* LT2 KDPGal aldolase were subjected to directed evolution by PCR mutagenesis and DNA shuffling, followed by DNA family shuffling of the most evolved *K. pneumoniae* (KP03-3) and *S. typhimurium* (ST04-5) KDPGal aldolase mutants with the most evolved *E. coli* KDPGal aldolase mutant (EC03-1).

The the *dgoA* mutants, EC03-1, KP03-3 and ST04-5, were subjected to DNA family shuffling using the single-stranded DNA shuffling method developed by Zhao and coworkers: Zha, W; Zhu, T; Zhao, H., *Methods Mol. Biol.* 2003, *231*, 91-97. Sequencing a small library of mutants (76) obtained from the family shuffling revealed a crossover rate of approximately 1.4 per gene using the published protocol. *E. coli* CB734 was transformed with a plasmid library that contained the chimeric *dgoA* hybrids (NR8.165), and the colonies that showed higher growth rate in the absence of aromatic amino acids supplementation as compared to *E. coli* CB734 carrying plasmid containing the parent gene, *EC03-1*, *KP03-3* or *ST04-5* (colonies appeared after 3 days on minimal salts medium) were selected. The *dgoA* gene hybrids from the five largest colonies that appeared after 2.5 days were sequenced and their encoding KDPGal aldolase activities toward DAHP formation were determined as shown in Table 9.

Table 9

Chimeric *dgoA* Genes Evolved by Cross-Species DNA Family Shuffling.

entry	family shuffling mutants	*dgoA* structure[a]	DAHP formation[b] (U[c]/mg)
1	NR8.165-2		1.31
2	NR8.165-3		0.30
3	NR8.165-4		0.22
4	NR8.165-5		0.10
5	NR8.165-6		0.56

[a] Symbol: EC03-1; KP03-3; ST04-5.

[b] Each *dgoA* mutant was inserted into the same plasmid (pJF118EH) with transcription controlled by a P_{tac} promoter. *E. coli* CB734 was used as host strain for expression of evolved enzymes. [c] One unit of DAHP synthase catalyzes the formation of one μmol of 3-dehydroshikimate per minute at 25°C.

[0076] All five mutants were found to be chimeras of genes from the *E. coli, K. pneumoniae and S. typhimurium.* Four of them contain two segments resulting from a single crossover event. One mutant, NR8.165-4, contains three segments resulting from two crossovers. It also noteworthy that most crossover events occured in the first 40-80 base pairs area where the three genes have 40-base pairs of nearly identical sequences. Compared with the wild-type *E. coli* KDPGal aldolase, the DgoA mutant NR8.165-4 has a 5-fold increase in k_{cat} and a 5-fold reduction in K_m for D-erythrose 4-phosphate, and thus a 25-fold increase in k_{cat}/K_m (entry 7, Table 10). Both the mutant NR8.165-2 and NR8.165-6 show a decrease in k_{cat}/K_m values relative to their parent enzymes (entries 6 and 8 vs. entries 4 and 5, Table 9). The K_m value ranging from 80 to 157 μM for D-erythrose 4-phosphate of the mutant enzymes is close to the K_m value of *E. coli* native DAHP synthase for D-erythrose 4-phosphate: AroF: K_m = 81.4 μM, k_{cat} = 29.5 s^{-1}; AroG: K_m = 141 μM, k_{cat} = 10.3 s^{-1}; AroH: K_m = 35 μM, k_{cat} = 20.6 s^{-1}. References are: Ramilo, C.A.; Evans, J.N.S., *Protein Express. Purif.* 1997, *9*, 253-261 (for AroF); Sheflyan, G.Y. et al., *J. Am. Chem. Soc.* 1998, *120*, 11027-11032 (for AroG); andAkowski, J.P.; Bauerle, R., *Biochemistry* 1997, *36*, 15817-15822

(for AroH). However, the k_{cat} of the mutant DgoAs is significantly lower than the k_{cat} of the native DAHP synthases.

Table 10

Kinetic Parameters of the Wild-Type KDPGal Aldolases and the Evolved Variants from Cross-Species DNA Family Shuffling.

Entry	DgoA[a]	K_m (E4P, μM)	k_{cat} (s^{-1})	k_{cat}/K_m (μM^{-1} s^{-1})
1	Wild-type *E. coli*	571	0.94	1.65×10^{-3}
2	wild-type *K. pneumoniae*	1507	1.39	3.22×10^{-4}
3	wild-type *S. typhimurium*	685	0.600	3.76×10^{-3}
4	EC03-1	124	2.49	2.01×10^{-2}
5	ST04-5	119	3.24	2.72×10^{-2}
6	NR8.165-2	157	2.51	1.60×10^{-2}
7	NR8.165-4	115	4.76	4.14×10^{-2}
8	NR8.165-6	80	0.504	6.30×10^{-3}

[a] All wild-type and the evolved KDPGal aldolases were expressed and purified as GST (glutathione S-transferase) fusion proteins.

Specific Example 5

Characterization of Pyruvate-Based Shikimate Synthesis Pathway In Vivo

[0077] To examine the functioning of created shikimate pathway variant in intact microbes, growth rates and synthesis of 3-dehydroshikimate were examined. *E. coli* CB734/pEC03-1 and *E. coli* CB734/pKP03-3 were completely dependent on plasmid-encoded, evolved DgoA isozymes EC03-1 and KPO3-3, respectively, for the formation of DAHP. *E. coli* CB734/pNR7.126 relied on plasmid-encoded, feedback-insensitive $AroF^{FBR}$ for DAHP synthase activity. When cultured under identical conditions where growth was dependent on de novo synthesis of aromatic amino acids and

aromatic vitamins, *E. coli* CB734/pEC03-1 and *E. coli* CB734/pKP03-3 entered the logarithmic phases of their growths 36 h and 12 h, respectively, later than *E. coli* CB734/pNR7.126 This is depicted in Figure 4, showing growth in the absence of aromatic amino acid and aromatic vitamin supplements in glucose-containing minimal salts medium under shake-flask conditions: *E. coli* CB734/pNR7.126 (squares); *E. coli* CB734/pEC03-1 (circles); *E. coli* CB734/pKP03-3 (triangles).

[0078] Synthesis of 3-dehydroshikimate employed *E. coli* NR7, which was constructed from *E. coli* KL3 using site-specific chromosomal insertions to inactivate all DAHP synthase isozymes. *E. coli* KL3 has been extensively used in studies examining the impact of phosphoenolpyruvate availability on the synthesis of 3-dehydroshikimate. Constructs were cultured under identical fermentor-controlled conditions. *E. coli* NR7/pKP03-3serA synthesized 8.3 g/L of 3-dehydroshikimate in 48 h in 5% yield from glucose. Only a trace amount of this product was synthesized by NR7/pNR8.074, which expressed plasmid-encoded, native *K. pneumoniae* DgoA. *E. coli* NR7/pEC03-1serA synthesized 12 g/L of 3-dehydroshikimate in 5% yield from glucose. For comparison, 2.0 g/L of 3-dehydroshikimate was synthesized in 0.9% yield by *E. coli* NR7/pNR8.075, which expressed plasmid-encoded, native *E. coli* DgoA.

[0079] Further characterization of the pyruvate-based shikimate pathways according to the present invention was performed in fed-batch fermentations (36 °C, pH 7.0, 20% air saturation, with growth on glucose-containing medium). In some cases, the host cell was further transformed with a polynucleotide encoding a transketolase (E.C. 2.2.1.1), a key enzyme responsible for the in vivo synthesis of E4P, whose low concentration can present a bottleneck in the DAHP synthesis process; overexpression thereof can enhance the yield of DAHP, and thus of DHQ, DHS, and further DHS derivates, e.g,. shikimate.. An *E. coli tktA* gene provided the coding sequence used therein (SEQ ID NO:9), although various transketolase isozymes may be used to supplement in vivo production of E4P, e.g., *tktB* gene-encoded enzymes, such as the *E. coli* TktB (SEQ ID NO:12). Alternatively, a transaldolase (E.C. 2.2.1.2) may be employed for this purpose. Examplary transaldolases include, e.g., E. coli isozymes TalA (e.g., GenBank Accession

No. D13159; gi:2337773) and TalB (e.g., GenBank Accession No. D13161; gi:2337775).

[0080] *E. coli* CB734 was not used for fed-batch fermentation of 3-dehydroshikimic acid in this study due to its L-leucine requirement and difficulty in comparing product titer and yield with previously reported 3-dehydroshikimic acid synthesis by constructs based on *E. coli* KL3. Therefore, instead of constructing a CB734*aroEydiB*⁻ strain, *E. coli* NR7 was constructed. All three DAHP synthase genes (*aroF, aroG, aroH*) in *E. coli* NR7 were inactivated by site-specific chromosomal insertions carried out in *E. coli* strain KL3 (AB2834 *serA::aroB*).

The DAHP synthase encoded by *aroF* and *aroH* in *E. coli* CB734 were inactivated by insertion with a chloramphenicol-resistant (Cm^R) gene and a kanamycin-resistance gene (Kan^R), respectively. P1-phage mediated transformation from *E. coli* CB734 could be the simplest way to disrupt the corresponding *aroF* and *aroH* genes in *E. coli* KL3 directly. Unfortunately, *E. coli* CB734 was found to be a P1 phage resistant strain possibly due to deletion of the *gal* operon in its chromosome. Transforming a plasmid-localized *galE* encoding UDP-galactose-4-epimerase in *E. coli* CB734 failed to reverse the P1 phage resistance phenotype of *E. coli* CB734.

Chromosomal inactivation of DAHP synthase genes *aroF*, *aroG*, *aroH* were then carried out by homologous recombination methods. Special recombinant-proficient *E. coli* hosts lacking exonuclease V of the RecBCD recombination complex are suitable for chromosomal recombination by transforming with linear DNA. Recombination can occur in *recB* or *recC* mutants carrying a suppressor *sbcB* mutation that enhances recombination by the RecF pathway or in *recD* mutants that are recombinase proficient but lack exonuclease V. A simple one-step method applicable to wild-type *E. coli* strain has been developed to use the bacteriophage λ Red recombinase to mediate recombination using linear DNA with short homolog extensions.

To construct *E. coli* NR7, the chloramphenicol-resistant (Cm^R) gene was inserted into *aroF* in a plasmid. The *aroF::*Cm^R allele was isolated and transformed into strain JC7623, a hyper-recombinant *recBC sbcBC* strain. Chloramphenicol-resistant transformants JC7623*aroF::*Cm^R in which the wild-

type *aroF* was exchanged with *aroF::Cm*^{R} allele by double-crossover event were obtained on chloramphenicol plates. P1-phage mediated transduction of JC7623*aroF::Cm*^{R} transferred the *aroF::Cm*^{R} mutation into KL3 to generate *E. coli* KL3*aroF::Cm*^{R}. Similarly, the *aroH::Kan*^{R} mutation was transferred from JC7623*aroH::Kan*^{R} to KL3*aroF::Cm*^{R} to prepare *E. coli* KL3*aroF::Cm*^{R} *aroH::Kan*^{R}. However, attempted transfer of the *aroG::tet* mutation by P1 phage mediated transduction from *JC7623aroG::Tc*^{R} was not successful. The *aroG* mutation was then generated using the λ Red recombinase method. An *aroG::Tc*^{R} DNA fragment was electroporated into KL3*aroF::Cm*^{R} *aroH::Kan*^{R} carrying a plasmid pKD46 encoding λ Red recombinase. Recombinants were selected for tetracycline resistance (5 μg/mL) at 30°C. Plasmid pKD46 was eliminated by growth at 42°C. Disruption of chromosomal *aroG* was confirmed by PCR from NR7 genomic DNA using a pair of primers flanked the *aroG* locus to amplify a fragment corresponding to the *aroG::tet* allele with correct size. The *E. coli* KL3*aroF::Cm*^{R} *aroG::Tc*^{R} *aroH::Kan*^{R} strain was designated as *E. coli* NR7.

In directed evolution of *E. coli* KDPGal aldolase, *E. coli dgoA* mutants were expressed under the control of a P_{trc} promoter in expression vector pTrc99A, while *K. pneumoniae* and *S. typhimurium dgoA* mutants were expressed under a *tac* promoter in pJF118EH. The *trc* promoter displays a spacing of 17 bp between the –35 and –10 consensus sequences[a] compared to a spacing of 16 bp between these regions in the *tac* promoter. Despite the 1 bp difference in spacing, P_{tac} and P_{trc} promoters are virtually of identical strength. However, plasmid pTrc99A does have a smaller size (4.2-kb vs. 5.3-kb in pJF118EH) and an increased plasmid copy number per chromosome (30 vs. 18 in pJF118EH) compared to plasmid pJF118EH. Therefore, the most active evolved *E. coli* mutant EC03-1 was excised from pEC03-1 and cloned into the pJF118EH vector to afford plasmid pNR8.140. Plasmid pNR8.158, pKP03-3serA and pST04-5serA were constructed by inserting a *serA* gene into the plasmids containing the corresponding *dgoA* mutants. Including the *serA* locus on plasmid provides the basis for plasmid maintenance during cultivation in minimal salts medium lacking L-serine supplementation. Furthermore, expression of the *dgoA* mutants in the same

plasmid enabled an unbiased comparison of the *in vivo* activities of the individually evolved KDPGal aldolases in terms of the production of the pyruvate-based shikimate pathway metabolite.

Table 11

Synthesis of 3-dehydroshikimic acid under fermentor-controlled conditions.

Entry	Construct	Genes	DHS[a] (g/L)	DHS (yield[b])
1	NR7/pKP03-3serA	*P_{tac}KP03-3, serA*	8.3	5.0%
2	NR7/pNR8.074	*P_{tac}wt-KPdgoA, serA*	0	0
3	NR7/pNR8.172	*P_{tac}EC03-1, serA*	5.1	2.4%
4	NR7/pNR8.170	*P_{tac}wt-ECdgoA, serA*	0	0
5	NR7/pST04-5serA	*P_{tac}ST04-5, serA*	6.9	3.3%
6	NR7/pNR8.121	*P_{tac}wt-STdgoA, serA*	0.1	0
7	NR7/pNR8.165-2serA	*P_{tac}NR8.165-2, serA*	7.4	3.3%
8	NR7/pNR8.165-4serA	*P_{tac}NR8.165-4, serA*	9.3	4.6%
9	NR7/pNR8.180	*P_{tac}NR8.165-4, serA, tktA*	12.4	6.0%
10	NR7/pNR8.182	*P_{tac}aroF^{FBR}, serA*	42.7	18%
11	NR7/pNR8.190	*P_{T5}NR8.165-4, serA, tktA*	10.5	6.5%

[a] DHS: 3-dehydroshikimic acid. [b] yield is calculated as (mol of DHS)/(mol of glucose).

Table 12

Evolved KDPGal aloldase activities towards DAHP formation.

Entry	Construct	DAHP formation assay (U[a]/mg)			
		12 h	24 h	36 h	48 h
1	NR7/pKP03-3serA	0	0.11	0.02	0.01

3	NR7/pNR8.172	0	0.05	0.05	0.05
5	NR7/ST04-5serA	0	0.30	0.25	0.22
8	NR7/pNR8.165-4serA	0	0.31	0.25	0.19
9	NR7/pNR8.180	0	0.13	0.17	0.15
11	NR7/pNR8.190	0	0.012	0.21	0.25

[a] One unit of DAHP synthase catalyzes the formation of one μmol of 3-dehydroshikimate per minute at 25°C. Isopropyl β-D-thioglucopyranoside (IPTG, 23.8 mg) was added at 12 h and every 6 h after.

E. coli NR7/pKP03-3serA was cultured under glucose-rich conditions at 36°C, 20% air saturation and pH 7.0 in a 2.0-L working volume fermentor, 8.3 g/L of 3-dehydroshikimic acid was produced after 48 h in 5% mol/mol yield from glucose (entry 1, Table 11). In contrast, only trace amount of 3-dehydroshikimic acid was observed in fermentation broth of NR7/pNR8.074 which encoded wild-type *K. pneumoniae dgoA* and *serA* genes (entry 2, Table 11). *E. coli* NR7/pNR8.172 produced 5.1 g/L of 3-dehydroshikimic acid in 2.4% mol/mol yield under the same conditions (entry 3, Table 11), while NR7/pNR8.170, which encoded wild-type *E. coli dgoA* and *serA* genes, produced only trace amount of 3-dehydroshikimic acid (entry 4, Table 11). *E. coli* NR7/pST04-5serA produced 7.1 g/L 3-dehydroshikimic acid in 3.4% yield (entry 5, Table 11). For comparison, NR7/pNR8.121 which encoded the wild-type *S. typhimurium dgoA* gene produced a trace amount of 3-dehydroshikimic acid (entry 6, Table 11).

The evolved KDPGal aldolase specific activities toward catalyzing the condensation of pyruvate and D-erythrose 4-phosphate were measured at 12, 24, 36 and 48 h after inoculation of the culture medium in the fed-batch fermentation runs (Table 12).

[0081] With DAHP formation catalyzed by an evolved KDPGal aldolase, the first reaction in the shikimate pathway can consume the pyruvate byproduct instead of competing for the phosphoenolpyruvate substrate required by PTS-mediated glucose transport. This constitutes a fundamental departure from all previous strategies employed to increase phosphoenolpyruvate availability in microbes such as *E. coli*.

[0082] The examples and other embodiments described herein are exemplary and not intended to be limiting in describing the full scope of compositions and methods of this invention. Equivalent changes, modifications and variations of specific embodiments, materials, compositions and methods may be made within the scope of the present invention, with substantially similar results.